We completely determine which matrices in $M_n(\mathbb{F})$ have an orthonormal eigenbasis.

### 4.4.1 Schur’s Lemma and the Spectral Theorem

Recall that a matrix $Q \in M_n(\mathbb{F})$ is orthonormal when $Q^H Q = I = QQ^H$.

**Definition 4.4.1.** Two matrices $A, B \in M_n(\mathbb{F})$ are orthonormally similar if there is an orthonormal matrix $U \in M_n(\mathbb{F})$ such that $B = U^H A U$.

**Note.** For an orthonormal matrix $U$ we have $U^{-1} = U^H$, and so

$$B = U^H A U = U^{-1} A U,$$

i.e., $A$ and $B$ are similar. The point here is that the inverse of an orthonormal matrix is easy to compute and we emphasize that in Definition 4.4.1.

Recall that a matrix $A \in M_n(\mathbb{F})$ is self-adjoint, or Hermitian, if $A^H = A$.

**Lemma 4.4.2.** If $A$ is Hermitian and orthonormally similar to $B$, then $B$ is Hermitian.

The proof of this is HW (Exercise 4.20).

Recall that the only matrices that are diagonalizable are the semisimple ones, i.e., the ones with an eigenbasis.

**Theorem 4.4.3 (Schur’s Lemma).** Every $A \in M_n(\mathbb{C})$ is orthonormally similar to an upper triangular matrix.

**Proof.** We prove Schur’s Lemma by induction.

The base case of $n = 1$ is trivial.

So suppose that the theorem holds for $n = k$, and take $A \in M_{k+1}(\mathbb{C})$.

Let $\lambda_1$ be an eigenvalue of $A$ and choose a unit eigenvector $w_1$.

Using the Gram-Schmidt algorithm, construct orthonormal vectors $\{w_2, \ldots, w_{k+1}\}$ so that $[w_1, w_2, \ldots, w_{k+1}]$ is an orthonormal basis for $\mathbb{C}^{k+1}$.

Let $U \in M_{k+1}(\mathbb{C})$ be the matrix whose $i^{th}$ column is $w_i$.

Then $U$ is an orthonormal matrix for which the product of the first row of $U^H$ with the first column of $AU$, i.e., $Aw_1$, is

$$w_1^H A w_1 = w_1^H (\lambda_1 w_1) = \lambda_1 w_1^H w_1 = \lambda_1.$$

The product the $i^{th}$ row of $U^H$ for $i = 2, \ldots, k + 1$ with $Aw_1$ is

$$w_i^H A w_1 = w_i^H (\lambda_1 w_1) = \lambda_1 w_i^H w_1 = 0$$

because the vectors $w_1, w_2, \ldots, w_{k+1}$ are orthonormal.

Thus the matrix $U^H A U$ has the block form

$$\begin{bmatrix} \lambda_1 & * \\ 0 & M \end{bmatrix}$$
where \(*\) is \(k\) row vector, \(0\) is a \(k\) column vector, and \(M \in M_k(\mathbb{C})\).

By the inductive hypothesis, there is an orthonormal matrix \(Q \in M_k(\mathbb{C})\) such that \(T_1 = Q_1^HMQ_1\) is an upper triangular matrix.

The matrix \(Q\) whose blocks are
\[
\begin{bmatrix}
1 & 0 \\
0 & Q_1
\end{bmatrix}
\]
is an orthonormal matrix for which
\[
(UQ)^H A(UQ) = Q^H(U^H A)Q = Q^H \begin{bmatrix}
\lambda_1 & * \\
0 & M
\end{bmatrix} Q = \begin{bmatrix}
\lambda_1 & * \\
0 & Q_1^HMQ_1
\end{bmatrix} = \begin{bmatrix}
\lambda_1 & * \\
0 & T_1
\end{bmatrix}.
\]

This is an upper triangular matrix and \(UQ\) is an orthonormal matrix.

Thus \(A\) is orthonormally similar to an upper triangular matrix. \(\square\)

**Remark 4.4.4.** An upper triangular matrix \(B\) that is orthonormally similar to a given matrix \(A\) is called a Schur form of \(A\). The similar matrices \(B\) and \(A\) have the same eigenvalues, and those eigenvalues have the same algebraic multiplicities and the same geometric multiplicities.

**Theorem 4.4.5.** Let \(\lambda\) be an eigenvalue of \(T \in \mathcal{L}(V)\) for a vector space \(V\) of dimension \(n\). If \(m_\lambda\) is the algebraic multiplicity of \(\lambda\), then
\[
\dim(\Sigma_\lambda(T)) \leq m_\lambda.
\]

**Proof.** Let \(v_1, \ldots, v_k\) be a basis for \(\Sigma_\lambda(T)\).

By the Extension Theorem, there are vectors \(v_{k+1}, \ldots, v_n\) so that \(S = [v_1, \ldots, v_n]\) is a basis for \(V\).

Since \(T(v_i) = \lambda v_i\) for all \(i = 1, \ldots, k\), the matrix representation of \(T\) in the basis \(S\) has the block form
\[
\begin{bmatrix}
\lambda I_k & * \\
0 & A_{22}
\end{bmatrix}
\]
where \(I_k\) is the \(k \times k\) identity matrix.

This block form implies that the characteristic polynomial \(p(z)\) of \(T\) factors as
\[
(z - \lambda)^k \det(zI_{n-k} - A_{22}).
\]

Since \((z - \lambda)^{m_\lambda}\) is a maximal factor of \(p(z)\), i.e., the power \(m_\lambda\) is the biggest possible, it follows that \(k \leq m_\lambda\). \(\square\)

**Corollary 4.4.6.** A matrix \(A \in M_n(\mathbb{C})\) is semisimple if and only if \(m_\lambda = \dim(\Sigma_\lambda(A))\) for every eigenvalue \(\lambda\) of \(A\).

**Spectral Theorem for Hermitian Matrices**

**Theorem 4.4.7 (First Spectral Theorem).** Every Hermitian matrix \(A \in M_n(\mathbb{C})\) is orthonormally diagonalizable, i.e., orthonormally similar to a diagonal matrix, and the diagonal matrix is real.
Proof. By Schur’s Lemma, $A$ is orthonormally similar to an upper triangular matrix $T$. Since $A$ is Hermitian and orthonormally similar to $T$, the matrix $T$ is also Hermitian by Lemma 4.4.2.

An upper triangular matrix that is Hermitian is diagonal, and because $T^H = T$, its diagonal entries are all real. □

Remark 4.4.8. The converse of Theorem 4.4.7 is also true: if $A$ is orthonormally similar to a real diagonal matrix, i.e., $A = U^H D U$ for $U$ orthonormal and $D$ a real diagonal matrix, then $A$ is Hermitian, i.e.,

$$A^H = (U^H D U)^H = U^H D^H U = U^H D U = A.$$ 

Corollary 4.4.9. If $A$ is Hermitian, then it has an orthonormal eigenbasis and the eigenvalues of $A$ are real.

Proof. Let $A$ be Hermitian.

By the First Spectral Theorem, there is an orthonormal matrix $U$ such that $U^H A U$ is a real diagonal matrix $D$.

Since $A$ and $D$ are similar, they have the same eigenvalues, so the eigenvalues of $A$ are real.

The columns of $U$ form an orthonormal eigenbasis for $A$. □

Remark 4.4.11. Not every eigenbasis of a Hermitian matrix is orthonormal. The eigenvectors need not be of unit length. More problematic is that for a eigenspace of dimension 2 or more, the eigenvectors found in this eigenspace need not be orthogonal.

Normal Matrices

Hermitian matrices are not the only matrices that are orthonormally diagonalizable.

Definition 4.4.12. A matrix $A \in M_n(\mathbb{C})$ is normal if $A^H A = A A^H$.


A skew-Hermitian matrix $B$ is normal because $B^H = -B$ implies

$$B^H B = -B^2 = BB^H.$$ 

An orthonormal matrix $U$ is normal because $U^H U = I = U U^H$ is $U^H U = U U^H$.

Theorem 4.4.14 (Second Spectral Theorem). A matrix $A \in M_n(\mathbb{C})$ is normal if and only if it is orthonormally diagonalizable.

Proof. Suppose $A$ is normal.

By Schur’s Lemma there is an orthonormal matrix $U$ and an upper triangular matrix $T$ such that $T = U^H A U$. Hence $T^H = U^H A^H U$ and

$$T^H T = U^H A^H U U^H A U = U^H A^H A U = U^H A A^H U$$

$$= U^H A U U^H A^H U = T T^H,$$

which shows that $T$ is normal.
For $T = [t_{ij}]$ where $t_{ij} = 0$ when $i > j$ (i.e., $T$ is upper triangular) we have

$$T^HT = \begin{bmatrix} t_{11} & 0 & \cdots & 0 \\ t_{12} & t_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ t_{1n} & t_{2n} & \cdots & t_{nn} \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{nn} \end{bmatrix},$$

and

$$TT^H = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{nn} \end{bmatrix} \begin{bmatrix} t_{11} & 0 & \cdots & 0 \\ t_{12} & t_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ t_{1n} & t_{2n} & \cdots & t_{nn} \end{bmatrix}.$$

Since $T$ is normal, the diagonal entries of $T^HT$ and $TT^H$ are the same:

$$|t_{11}|^2 = |t_{11}|^2 + |t_{12}|^2 + \cdots + |t_{1n}|^2$$

$$|t_{12}|^2 + |t_{22}|^2 = |t_{22}|^2 + \cdots + |t_{2n}|^2$$

$$\vdots$$

$$|t_{1n}|^2 + |t_{2n}|^2 + \cdots + |t_{nn}|^2 = |t_{nn}|^2.$$

These imply that $t_{ij} = 0$ whenever $i \neq j$, and so $T$ is diagonal.

Now suppose that $A$ is orthonormally diagonalizable: there is an orthonormal matrix $U$ and a diagonal matrix $D$ such that $D = U^HAU$.

Then $A = UD^HU^H$ and $A^H = UD^HUU^H$.

Because $D$ is diagonal we have $D^HD = DD^H$ (i.e., $D$ is normal).

Hence

$$A^HA = UD^HU^HU^DU = UDD^HUU^H = U^HD^HU^HU^H = UD^HUU^HUDU^H = AA^H.$$

Thus $A$ is normal. \qed